- A3273-8
1/11/88

Vol. 63 No. 1 January 1988

JOURNAL OF Medical Education

PUBLICATION OF THE ASSOCIATION OF AMERICAN MEDICAL COLLEGES

Combined B.S.-M.D. Programs

Comparison of Performances of Students in Programs at LSU Medical School in Shreveport
F. Scott Kennedy and Judy C. Austin

Evaluation of Gannon-Hahnemann Program To Provide Family Physicians for Underserved Area
Frederick J. Bloom, Mark K. Vatavuk, and Susan E. Kaliszewski

Student Evaluation of Accelerated Program at the University of Miami
Jeffrey P. Jacobs, Robert E. Hinkley, and J. Phillip Pennell

Time for a Change? (Editorial)
G. William Bates

An Interim Report on the MCAT Essay Pilot Project
Judith Anderson Koenig and Karen J. Mitchell

Types of Medical Writing and Teaching of Writing in U.S. Medical Schools
Karin L. Yanoff and Fredric D. Burg

Journal of Medical Education

Volume 63 • Number 1 • January 1988

Combined B.S.-M.D. Programs

Comparison of Performances of Students in Programs at LSU Medical School in Shreveport—F. Scott Kennedy and Judy C. Austin 1

Evaluation of Gannon-Hahnemann Program To Provide Family Physicians for Underserved Area—Frederick J. Bloom, Mark K. Vatavuk, and Susan E. Kaliszewski 7

Student Evaluation of Accelerated Program at the University of Miami—Jeffrey P. Jacobs, Robert E. Hinkley, and J. Phillip Pennell 11

Time for a Change? (Editorial)—G. William Bates 19

An Interim Report on the MCAT Essay Pilot Project—Judith Anderson Koenig and Karen J. Mitchell 21

Types of Medical Writing and Teaching of Writing in U.S. Medical Schools—Karin L. Yanoff and Fredric D. Burg 30

Medical Students' Access to Information and Resources for the Residency Selection Process—Mary P. Taggart, Steven A. Wartman, and Albert F. Wessen 38

Patterns of Psychoactive Drug Prescriptions by House Officers for Nonpatients—Allan W. Clark, Jerald Kay, and David C. Clark 44

A Program To Teach House Calls for the Elderly to Fourth-Year Medical Students—Ann E. K. Page, Leslie Walker-Bartnick, George A. Taler, Dorothy A. Snow, Debra S. Wertheimer, and Mohamed S. Al-Ibrahim 51

Communications: Basic Science Course Grades as Predictors of Performance on the NBME Part I Examination—John Yergan, Douglas C. Schaad, Jan D. Carline, Theodore J. Phillips, D. Daniel Hunt, and Daniel O. Graney 59

The Effect of Human Dissection on First-Year Students and Implications for the Doctor-Patient Relationship—Norman Gustavson 62

Effect of Required Reading in a Surgical Clerkship on Performance on the NBME Surgery Examination—Gregory V. Stanislav and Alan H. Fruin 64

Briefs 66

Book Review 69

Bibliography 70

Letters to the Editor 76

Officers of the Association and Members of the Executive Council, 1987–88

CHAIRMAN: John W. Colloton, *University of Iowa Hospitals and Clinics*
CHAIRMAN-ELECT: D. Kay Clawson, *University of Kansas*
IMMEDIATE PAST CHAIRMAN: Edward J. Stemmler, *University of Pennsylvania*
PRESIDENT: Robert G. Petersdorf, *AAMC, Washington, D.C.*

COUNCIL OF DEANS
William T. Butler, *Baylor College of Medicine**
L. Thompson Bowles, *George Washington University*
William B. Deal, *University of Florida*
John Naughton, *SUNY at Buffalo*
Richard S. Ross, *Johns Hopkins University*
Henry P. Russe, *Rush Medical College*
Robert Tranquada, *University of Southern California*
W. Donald Weston, *Michigan State University*

COUNCIL OF ACADEMIC SOCIETIES
Douglas E. Kelly, *University of Southern California**
William F. Ganong, *UC at San Francisco*
Ernst Jaffe, *Albert Einstein College of Medicine*
Frank G. Moody, *University of Texas, Houston*

COUNCIL OF TEACHING HOSPITALS
J. Robert Buchanan, *Massachusetts General Hospital**
Spencer Foreman, *Montefiore Medical Center*
Gary Gambuti, *St. Luke's-Roosevelt Hospital Center*
James Mongan, *Truman Medical Center*

ORGANIZATION OF STUDENT REPRESENTATIVES
Kimberly Dunn, *University of Texas, Houston**
Clayton Ballantine, *University of Louisville*

DISTINGUISHED SERVICE MEMBER
Edward N. Brandt, Jr., *University of Maryland*

* Chairman of respective organization.

JOURNAL OF Medical Education

Publication of the Association of American Medical Colleges

Editorial Office and Staff

Association of American Medical Colleges
One Dupont Circle, N.W., Washington, D.C. 20036 (202) 828-0660

Editor: Merrill T. McCord
Associate Editor: James R. Ingram
Assistant Editor: Addeane S. Caelleigh

Editorial Assistants: C. Trudel Turner
Vickie L. Wilson

Information for Contributors

Features. The *Journal of Medical Education* considers four types of papers from contributors: major articles, communications, briefs, and letters to the editor. *Major articles* include reflective, perspective writings and fully documented reports on major studies involving and affecting medical education. *Communications* and *briefs* are for presentations of new ideas on medical education, preliminary findings of projects, and reports on limited studies. *Letters* can be on any subject related to medical education.

Requirements. Major articles are limited to 15 double-spaced typed pages of text, 15 references (30 for literature reviews), and five tables or figures and must be accompanied by an abstract of no more than 150 words. Communications are limited to four double-spaced typed pages of text with no tables or figures or to three double-spaced typed pages of text with one table or figure; references in communications are limited to 10. Briefs are limited to 250 words and do not carry tables, figures, footnotes, or references. Letters to the editor are limited to two double-spaced typed pages of text with no tables, figures, footnotes, or references (cited publications can be identified in text). All copy, including references, tables, and legends, should be typed double-spaced on one side of letter-sized white paper with standard margins. Pages should be numbered.

All manuscripts, except letters, should be written in the third person, and all should be submitted in duplicate. Poor reproduction copies will not be considered. Manuscript titles must not run more than two *Journal* lines, and all articles should contain subheads of two or three words each. Footnotes should be kept to a minimum and should be no more than two or three lines in length. All authors should be fully identified in one succinct paragraph. Only doctoral degrees should be listed.

Each table, diagram, figure, and photograph should have a legend and should be understandable independent of the text. Each table should be typed on a separate sheet. Simple tables should be written into text in narrative form. Vertical rules are not used in tables; only straddle horizontal rules joining two or more columns are permissible. All illustrations must be professionally drawn and lettered, and the lettering should be in a type large enough that it will be legible when the illustration is reduced to fit a *Journal* page. Either the original artwork or a good, high-contrast photograph of the artwork should be submitted. Identify each figure by number.

Only formally published material specifically cited in the text should be included in the reference section. References should be listed in the order in which they are cited in the text and must be submitted in accordance with the style followed in current issues of the *Journal*, including, in order: author, title, journal abbreviation (*Index Medicus* form), volume number, inclusive page numbers, and year; book references should also include editors, edition, publisher, and place of publication. *A detailed style sheet of instructions for listing references is available from the editor.*

General Information. Manuscripts are considered with the understanding that they are not under consideration elsewhere and have not been previously published in any form.

Personal acknowledgments are not published, and the number of listed authors is limited to six. The first dictionary (Webster's) spelling for all words is used. Italics for emphasis should be avoided. Outline presentations are not acceptable.

Authors should study a few recent issues of the *Journal* for additional guidelines on style and format. Special requirements apply to supplements, and the editor should be queried before lengthy material is submitted.

Because of the volume of manuscripts received by the *Journal* and because less than half of the contributions can be accepted for publication, no more than one paper by a given author will be considered at a time, and no more than one paper on a given project will be considered. Two-part articles also should not be submitted.

All contributions are reviewed by the Editorial Board for acceptability. A list of the elements considered by the reviewers appears in the September issue. All manuscripts accepted for publication are subject to editing and condensation by the editorial staff. Edited copy is sent to the author. Galley proofs also are mailed to authors for checking and should be returned within 24 hours after receipt.

Reprints. Reprint order forms are sent to the author at the same time as galley proofs. Reprints must be ordered at this time and are not available at a later date. The *Journal* editorial office does not maintain a file of reprints. Reprints must be ordered in multiples of 100; prices depend on the length of the article.

Correspondence

Address correspondence regarding manuscripts and advertising to the Editor, Association of American Medical Colleges, One Dupont Circle, N.W., Washington, D.C. 20036. Address correspondence concerning subscriptions to Membership and Publication Orders in care of the Association. Address requests for reprints to authors of articles, *not* to the *Journal*. POSTMASTER: Send address changes to: *Journal of Medical Education*, AAMC, Suite 200, One Dupont Circle, N.W., Washington, D.C. 20036.

The *Journal of Medical Education* (ISSN 0022-2577) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, N.W., Washington, D.C. 20036, and printed at Waverly Press, Inc., Mt. Royal and Guilford Aves., Baltimore, Maryland 21202. Second-class postage paid at Washington, D.C., and additional mailing offices.

Subscription Rates: $30.00 per year, $5.00 per single copy; students, $25.00 per year; foreign, $40.00 per year, $6.00 per single copy; Latin America and Canada, $30.00 per year.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

Copyright 1988 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. **Printed in U.S.A.**

An Interim Report on the MCAT Essay Pilot Project

Judith Anderson Koenig and Karen J. Mitchell, Ph.D.

Abstract—Results from four pilot administrations of the Medical College Admission Test essay question are reported. Analyses focused on (*a*) the performance characteristics of sample groups differentiated by gender, size of hometown, race/ethnicity, and dominant language; (*b*) the relationships between essay scores and academic/demographic characteristics; and (*c*) the reliability of one 45-minute versus two 30-minute essays. No differences were found for examinees grouped by gender and size of home community. Mean differences among the racial/ethnic groups were explained largely by reading level differences. Differences in essay performance by language group were large and unexplained by reading level differences. No relationship was found between the essay score and the academic/demographic characteristics. Reliability estimates for two 30-minute essays were higher than for one 45-minute essay; however, the 30-minute period yielded writing of poorer quality. Test-retest reliabilities for the 45-minute topics will remain the focus of future studies as will performance by examinees for whom English is a second language. The impact of the essay on the selection process will also be assessed.

The Association of American Medical Colleges is investigating the desirability and feasibility of including an essay as part of the Medical College Admission Test (MCAT). This endeavor, entitled the MCAT Essay Pilot Project, called for the administration of an essay on a trial basis with each of the spring and fall MCAT administrations in 1985, 1986, and 1987. The overall objectives of the project are to plan, develop, implement, and evaluate an essay written by MCAT examinees under standard conditions and in response to a topic developed with specific criteria.

Ms. Koenig is a research associate, and Dr. Mitchell is director, Medical College Admission Test Program, Section for Student and Educational Programs, Association of American Medical Colleges, Washington, D.C.

The MCAT essay was administered in the spring and fall of 1985 and of 1986. The essay topics were designed to assess an examinee's skill in: developing a central idea; synthesizing concepts and ideas; separating relevant from irrelevant information; developing alternative hypotheses; presenting ideas cohesively and logically; and writing clearly and observing the accepted practices of grammar, syntax, punctuation, and spelling consistent with timed, first-draft composition.

Under the guidance of the Ad Hoc Advisory Committee to the MCAT Essay Pilot Project, an evaluation plan was developed to determine whether the essay should become a permanent part of the MCAT. The evaluation plan is divided into four phases, each containing two primary questions:

Phase 1—What is the nature of the information provided by an essay? What are the performance characteristics of various examinee groups?

Phase 2—What is the impact of the essay on the selection process? Is the information provided by the essay unique and useful in selecting students?

Phase 3—What effect does an essay on the MCAT have on undergraduate students' course selections? Does the presence of an essay on the MCAT affect the undergraduate curriculum or the types of applicants?

Phase 4—What are the costs associated with the development, administration, and distribution of an MCAT essay? What different methods (and their costs) are available for the evaluation and distribution of essays?

The data reported below provide preliminary information on phase 1 and part of phase 2 of the evaluation plan. The following research questions were addressed for each administration:

1. What are the performance characteristics of the total sample and of sample groups differentiated by gender, size of hometown, race, and dominant language?
2. What are the relationships between essay scores and such demographic/academic characteristics as age, years of postsecondary education, and selectivity of undergraduate college attended?
3. What are the relationships between essay performance and scores on the MCAT science and skills analysis tests?
4. What level of reliability can be achieved with an essay question?

Method

With each MCAT administration in 1985 and 1986, an essay test was included. A 45-minute period was allotted for the essay during the first three administrations. The essay was expanded to 60 minutes in the fall of 1986. A different topic was used for each administration.

For each administration, a sample of examinees was selected to represent the demographic and academic characteristics of the MCAT examinee population as a whole. Papers for both initial and repeat examinees who indicated a preference for allopathic medicine were included in the samples. The sample essays were scored by readers from the California university system who were experienced in scoring for large-scale writing projects.

The essays were scored using a holistic approach that regards a piece of writing as a unit without subscores or separable aspects. A score is assigned based on a general impression of the overall quality of the entire paper using a scoring guide that explicates the characteristics of essays at each score level. The raters are trained to internalize the scoring guide and to minimize the effects of individual biases. The actual scoring does not begin until all readers are in close agreement during practice scorings of sample papers. Calibration exercises are performed prior to the actual scoring and approximately every two hours during the scoring to ascertain that all readers remain correctly focused in using the scoring guide.

The scoring room is arranged so that six to eight readers sit at one table. Assigned to each table is a "table leader." Table leaders are selected based on their experience with holistic scoring. Their job is to provide guidance to the scorers in their determination of scores. The rater training and scoring sessions were structured using White's (1) suggestions for organizing and managing holistic essay readings.

Essays were randomly sorted into groups of 20, and each batch was assigned at random to two readers. Each essay was

rated on a 6-point scale by two readers; each reader assigned ratings without knowledge of the other rating. The two ratings on each essay were subsequently checked for score agreement. Papers receiving ratings that were more than a score point apart were given a third reading by one of the table leaders. In the instances of two readers, the final score was the sum of the two scores, resulting in a score ranging from 2 to 12. When a third reading was needed, the final score was determined by the table leader and was usually the sum of the third score and one of the first two scores.

The first three MCAT essay administrations contained one essay question. In the fourth administration, in the fall of 1986, all examinees were asked to write on two essay topics. The scoring procedure for this administration was identical to previous scorings except that the final scores for each topic were averaged to produce a score consistent with scores in the previous administrations. Again, the score scale ranged from 2 to 12.

Means and standard deviations for essay scores were calculated separately for students grouped by gender, race/ethnicity, size of hometown (less than or more than 10,000 population), and dominant language for each administration. Examinees were grouped into four racial/ethnic groups by their self-identification as white, black, Asian/Pacific Islander, or Hispanic. Those who described themselves as Mexican American, mainland Puerto Rican, or other Hispanic were included in the Hispanic category. Since the number of Native American examinees was so small, data for this group were not reported. The two language groups consisted of commonwealth Puerto Ricans and all others. Because the MCAT does not ask examinees to indicate their native language, it was thought that the commonwealth Puerto Rican group would be likely to contain large numbers who spoke English as a second language. Since preliminary examination of the data revealed differences by racial/ethnic group and by language group (that is, between commonwealth Puerto Ricans and all others), average essay scores were computed for students divided into groups by their scores on the MCAT skills analysis: reading subtest to determine whether the differences in essay scores observed between racial/ethnic and language groups were attributable to differences in average reading levels. An analysis of covariance was performed with race/ethnicity as the independent variable and the reading subtest score as the covariate. The reading score was chosen as a covariate because the reading subtest and the essay question are designed to assess some of the same skills and because the reading score was the most highly correlated of all subtests with the essay.

Means and standard deviations for essay data were also computed for examinees grouped by number of semester hours of English, undergraduate major, self-ratings of writing and reading ability, and the selectivity of the undergraduate college. Self-rating of reading and writing was assessed by responses to questions in the MCAT registration packet asking for a comparison of the student's own ability with that of other college students. A 5-point scale was used consisting of the following choices: below average, average, above average, top 10 percent, and top 1 percent. Pearson's product-moment correlations were used to determine the relationships between essay scores and these variables.

The relationships between the six MCAT subtests and the essay were also examined using Pearson's product-moment correlations. Simultaneous multiple

regression analysis was employed to determine the degree to which the essay score could be predicted from the current six-test battery. Reliability estimates were obtained using generalizability theory techniques (2, 3).

Results for Spring 1985–Spring 1986

Essay results for the spring 1985, fall 1985, and spring 1986 examinees (who wrote one 45-minute essay) appear in Figure 1. (The results for the fall 1986 sample are discussed in the next section.) The mean score for the spring 1985 sample of 3,118 was 6.8; for the fall 1985 sample of 1,639 it was 7.6; and for the spring 1986 sample of 4,074 it was 7.1. The standard deviations for the three testings ranged from 1.6 to 1.7. In each case, the data were normally distributed, and all score points were represented.

Group data for all three samples are presented in Table 1. Group differences were negligible for gender and size of hometown but were somewhat larger for race/ethnicity and dominant language. The rank orders of the groups were similar across the three administrations.

Results from the analyses of covariance indicated that a significant amount of the variation in essay performance among the racial/ethnic groups was due to differences in reading level. Analysis of covariance also allows comparison of group means with the overall mean while controlling for reading level. Examination of the means adjusted for reading level revealed that the differences between the adjusted group mean and the adjusted overall mean ranged from −.25 to .20 of a score point for blacks (n = 173, 163, 263, respectively for the three administrations), .09 to .30 for whites (n = 2244, 991, 3051), −.02 to −.38 for Hispanics (n

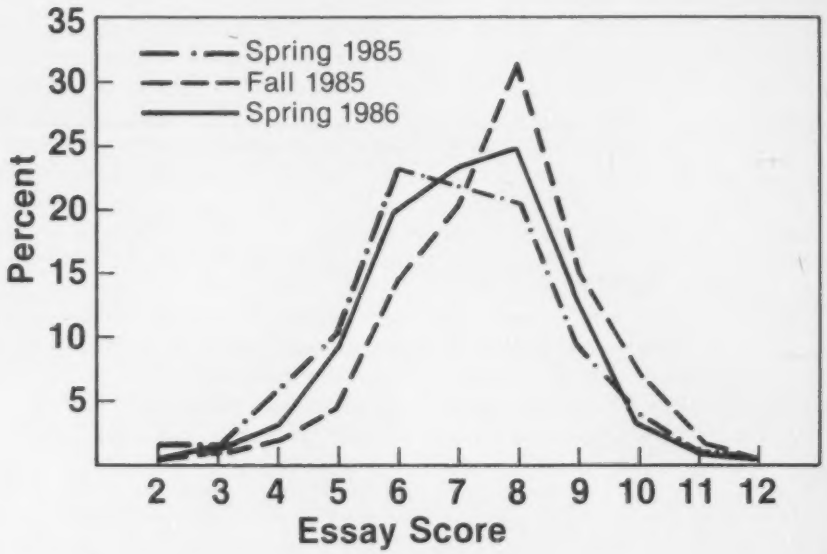

Figure 1
Percentages of representative sample groups of examinees (spring 1985 = 3,118; fall 1985 = 1,639; and spring 1986 = 4,074) at each score level on the essay for three administrations of the Medical College Admission Test.

TABLE 1

Numbers of Sample Group Examinees and Their Essay Scores for Three Administrations of the Medical College Admission Test, by Four Characteristics*

Characteristics	Spring 1985			Fall 1985			Spring 1986		
	Mean Score	S.D.	No.	Mean Score	S.D.	No.	Mean Score	S.D.	No.
Gender									
Male	6.8	1.7	1,971	7.5	1.5	934	7.0	1.6	2,544
Female	7.0	1.7	1,147	7.7	1.6	705	7.3	1.5	1,530
Hometown size†									
Less than 10,000	6.8	1.7	474	7.6	1.6	246	7.0	1.5	689
More than 10,000	6.8	1.8	2,336	7.6	1.6	1,134	7.1	1.5	2,981
Race/ethnicity‡									
Black	5.9	1.6	173	6.9	1.5	163	6.4	1.6	263
White	7.1	1.6	2,244	7.9	1.4	991	7.3	1.4	3,051
Hispanic	6.5	1.7	348	6.9	1.7	166	6.6	1.5	152
Asian	6.6	1.9	269	7.6	1.7	146	6.9	1.7	392
Language§									
Commonwealth Puerto Rican	3.7	1.6	66	4.4	1.5	158	4.1	1.6	188
Others	6.8	1.7	3,052	7.6	1.5	1,481	7.1	1.5	3,886

* The sample groups were representative of the examinee pool as a whole; both first time and repeat examinees who indicated a preference for allopathic medicine were included. The sample sizes were: spring 1985 = 3,118; fall 1985 = 1,639; and spring 1986 = 4,074.

† Numbers do not add to the total sample size because of incomplete examinee records.

‡ Race/ethnicity was based upon self-identification on the MCAT. Hispanics included Mexican Americans, mainland Puerto Ricans, and other Hispanics. Not all races are listed.

§ The commonwealth Puerto Rican examinees were considered to speak English as a second language; all others were considered to be native English speakers.

= 348, 166, 152), and −.10 to .07 for Asians (n = 269, 146, 392). Hence, even though there were differences in essay performance by racial/ethnic group, these differences decreased when basic skills or reading levels were considered. However, the commonwealth Puerto Ricans, (n = 66, 158, 188) scored 1.75 to 2.00 points below the adjusted overall mean essay score. Factors other than reading level differences may have contributed to these examinees' lower performance.

The type and magnitude of differences among racial/ethnic groups were similar for all three administrations. However, the mean essay score for all racial/ethnic groups in the spring 1986 sample improved slightly over prior administrations when reading level was held constant. This improvement may have resulted from the availability of practice materials and topics distributed from previous administrations.

No relationship was found between essay performance and age, years of postsecondary education, and number of semester hours in English courses. The relationship between essay performance and number of English semester hours may have been obscured by an inability to identify course hours for remediation or reinforcement and those for advanced English courses. The humanities and social science majors had higher mean scores than examinees with other majors. There was a positive relationship between essay scores and examinees' self-ratings in writing and reading. The examinees, thus, proved to be good judges of their writing ability. There was also a positive relation-

ship between essay performance and college selectivity. Students from selective undergraduate institutions received higher essay scores, and those from less selective schools received lower scores. Selectivity was defined by the average combined Scholastic Aptitude Test score for all students at an institution, as discussed by Astin (4). Astin index data for 1977 were used in the present study, as it was the most recent available.

Correlations between the essay and other MCAT subtests are shown in Table 2. Correlations between the essay and science areas ranged from .25 to .30, whereas those between the essay and skills analysis tests were higher. The reading subtest had the highest correlations with the essay, with coefficients that ranged from .42 to .49. The correlations in Table 2, however, were lower than those observed among the six MCAT subtests themselves; those correlations ranged from .51 to .87. These correlations indicate that the essay was measuring skills different from those assessed by the MCAT six-test battery.

When essay scores were predicted from data for the six MCAT subtests, the multiple correlations ranged from .44 to .49 for the three essay administrations. That is, from 19 to 25 percent of the variance in the essay score distribution overlapped with variance in the subtests. It is worth noting that the simple correlation between the reading subtest score and the essay score was of the same magnitude as the multiple correlation between the six subtest scores and the essay score. This similarity was due to the high intercorrelation among the four MCAT science subtests and their low correlation with the essay. When these four scores are considered jointly with the two skills analysis scores, very little unique information is added, and the multiple correlation remains in the same range as the simple correlation.

The interrater reliabilities for all three essay scorings ranged from .72 to .73. These estimates indicate the amount of variation in the essay scores that was attributable to the differences in the examinees' writing ability rather than to such sources of error as differences among readers or time or day of scoring (5). The reliability estimates do not address differences in performance due to variation in

TABLE 2

Correlations Between Essay Scores and Subtest Scores of Sample Groups of Examinees for Three Administrations of the Medical College Admission Test*

Subtests	Essay		
	Spring 1985	Fall 1985	Spring 1986
Biology	.29	.30	.27
Chemistry	.28	.27	.25
Physics	.27	.26	.26
Science problems	.29	.29	.30
Skills analysis: reading	.43	.49	.42
Skills analysis: quantitative	.38	.37	.36
All subtests	.45	.49	.44

* The sample groups were representative of the examinee pool as a whole; first-time and repeat examinees indicating a preference for allopathic medicine were included.

the content or skill requirements of different topics.

Results for Fall 1986

Results from the three administrations presented above suggested that additional research was needed on the development of essay topics that are comparable in difficulty and in assessing prerequisite skills. Test-retest reliabilities for small samples of examinees and differences among the mean essay scores for the three administrations indicated that changes in the format or focus of essay questions were warranted. Testing of four essay topics differing in focus and length of response was undertaken with the fall 1986 MCAT administration. Each topic was paired once with every other topic and administered in a 30-minute time period. Counterbalancing was used so that each topic appeared as the first and the second essay. This produced 12 different topic combinations. The examinees wrote two different essays, and the essay testing time was increased to 60 minutes. Each topic was also administered in a 45-minute time period. Two of the four topics were shortened to be 15-minute essays. These were paired with the 45-minute essays so that testing time would also be 60 minutes. In this case, the 45-minute essay was always administered first. Random samples of testing centers were selected for the 30/30- or 45/15-minute forms. It was hypothesized that the administration of two topics would lead to a better evaluation of the skills measured by essay writing and, hence, would result in increased reliability and comparability of essay data. Analyses were planned to answer two questions: (*a*) Could 45-minute questions be developed that were similar in difficulty and in skills assessed across administrations? (*b*) Would the administration of a 60-minute essay component with two 30-minute topics lead to sufficient comparability and reliability of essay data?

The fall 1986 data revealed that the same topic administered in a 45- and 30-minute time period elicited a wide range of response. Essays written in the 15-minute period were not scored. Mean scores for the 45-minute topics ranged from 7.1 to 7.7. Test-retest reliabilities were calculated for each of the four 45-minute topics on samples of repeating examinees who sat for both the spring and fall administrations in 1986 ($n = 139$ to 179). The four reliability estimates ranged from .36 to .58.

Mean scores for the 60-minute format ranged from 7.1 to 7.4, with reliability coefficients ranging from .60 to .81. These reliability estimates were based on split-half and generalizability coefficients. Test-retest reliability estimates were not available for the 60-minute format because this test format was not administered prior to fall 1986. However, test-retest reliabilities can be extrapolated from these estimates and would generally be expected to exceed those of the 45-minute format and fall in the low 70s.

Discussion

The statistical characteristics of the 1985 and 1986 essays considered by administration were encouraging. No average score differences were evident between examinees grouped by the various academic/demographic characteristics. There were differences for examinees grouped by racial/ethnic status, but these differences were largely related to reading level. Larger differences in mean scores were found between examinees grouped by language (commonwealth Puerto Ricans and all others), and these differences were not explained by the variance in reading level. The correlations between the essay

and the other MCAT subtests indicated that the essay was assessing a skill or skills unexamined by the other tests.

The interrater reliabilities were at acceptable levels. Test-retest reliabilities for the 45-minute topics were low, while the split-half reliabilities between the two 30-minute essays within the 60-minute format administered in fall 1986 were encouraging. These split-half reliabilities were sufficiently high to expect test-retest reliabilities for 60-minute formats to reach acceptable levels. However, the 45- and 30-minute topics elicited papers of notably different quality. The 45-minute period seemed to allow for fuller treatment of issues and more complete argument development. The skills involved in organizing and presenting ideas and sustaining a line of reasoning were demonstrated better in a 45-minute than in a 30-minute period. The 30-minute topics produced papers with more simplistic arguments and increased numbers of surface errors (such as, punctuation and word usage). The quality differences between 45- and 30-minute topics were not reflected in mean score discrepancies for the two types of questions because each was graded against time-specific criteria.

Conclusions

The 45-minute period seemed to allow examinees to plan, organize, and execute a thoughtful essay that would be a useful addition to the application process. Concerns about the quality of writing in the 30-minute essays were too marked to promote the use of the 30-minute period for administration of an essay. Future research will focus on methods for increasing the reliability of the 45-minute topics.

The spring 1987 MCAT contained several additional questions about the examinee's native and/or second language. This information will be used to investigate more thoroughly the differences in essay performance for special groups of examinees.

Future work will also focus on answering questions presented in phases 2 through 4 of the evaluation plan. Validity data are being collected for a small number of students currently enrolled in medical school who have agreed to write essays. The relationship between their essay scores and medical school performance will be examined, and results should be available in 1988. The impact of the essay on the selection process will be investigated by schools participating in simulated admissions decision-making exercises and retrospective selection exercises. Several schools have already completed these exercises, and the results will be reported. Essays will be used by pilot schools in a variety of ways for selection of the 1988 entering class. These schools will use essays at differing points in their admissions process. Some will use them as an initial screening device, while others will consider them later in the admissions process, such as in conjunction with the interview.

Research on the impact of the essay on undergraduate students' course selection, such as the addition of writing courses, and on the types of medical school applicants, such as an increase in the number of applicants with nonscience majors, has been designed. Cost data on the development, administration, and distribution of the essay are being compiled.

References

1. WHITE, E. *Teaching and Assessing Writing: Recent Advances in Understanding, Evaluating, and Improving Student Performance.* San Francisco, California: Jossey-Bass Publishers, 1985.
2. BRENNAN, R. L. *Elements of Generalizability Theory.* Iowa City, Iowa: American College Testing Program, 1983.

3. CARDINET, J., TOURNEUR, Y., and ALLAL, L. The Symmetry of Generalizability Theory: Applications to Educational Measurement. *J. Educ. Meas.*, **13**:119–135, 1976.
4. ASTIN, A. W. *Who Goes Where to College?* Chicago, Illinois: Science Research Associates, 1965.
5. MITCHELL, K., and ANDERSON, J. Reliability of Holistic Scoring for the MCAT Essay. *Educ. Psychol. Meas.*, **46**:771–775, 1986.

Types of Medical Writing and Teaching of Writing In U.S. Medical Schools

Karin L. Yanoff and Fredric D. Burg, M.D.

Abstract—U.S. medical schools were surveyed to determine which types of medical writing are most important for physicians, house officers, and medical students to learn and whether such types are formally taught. According to the responding schools, the five most important types were: write-up of the patient history and physical examination, progress note and discharge summary (tied), peer-reviewed published paper (of either clinical or laboratory research), and grant proposal. Certain types, however, were considered more important for some categories of physicians and trainees than others. Although these major types of writing represent complex and diverse writing tasks, they are not taught extensively in medical schools or thoroughly explained in the literature on medical writing.

The quality of medical writing has been a longtime concern (1–4), and the negative consequences of ineffective writing by physicians have also been documented, for instance, failure to publish research, loss of grant proposals, increased litigation, and reduced patient compliance (5). Little research, however, has focused on the kinds of writing physicians actually do and need to learn to do well.

In a 1982 survey of U.S. medical school deans, Bjork and Oye (5) found that 15 of the 101 schools responding offered formal writing courses as electives and recommended that writing instruction be made part of the medical school curriculum. However, these authors did not determine what types of writing were, or should be, taught. Preliminary to designing a writing program for the University of Pennsylvania School of Medicine, the authors of the present report attempted to discover which kinds of writing are most important for physicians to learn and whether these types are taught.

Methods

For the purposes of the survey reported here, medical writing was defined as those texts written, or sometimes dictated and edited, for a variety of purposes and audiences by medical students and house officers and by physicians in practice.

To determine the specific types of medical writing and the emphasis given to each, the authors examined two types of literature: first, 33 general texts on medical writing published between 1900 and 1985 and, second, books and articles written since 1966 on specific types of medi-

Ms. Yanoff is associate professor of English, Community College of Philadelphia, and Dr. Burg is associate dean for academic programs, University of Pennsylvania School of Medicine, and professor of pediatrics, Children's Hospital of Philadelphia.

cal writing. In this examination, the frequency of reference to a type of medical writing was assumed to indicate the relative importance of that type to physician-writers. The types of writing most frequently referred to in the general texts on medical writing were: scientific paper (that is, peer-reviewed published paper, of either clinical or laboratory research, other than the case report), case report, and abstract. Those types most frequently referred to in the books and articles on specific types of medical writing were: scientific paper, grant proposal, and write-up of history and physical examination.

These types of writing identified in the literature were discussed in interviews with eight individuals at different levels of medical training and in different settings, with four administrators concerned with medical education at the University of Pennsylvania School of Medicine, and with three editors of national medical journals. In considering the types of writing done in the practice of medicine, the authors categorized writers according to their level of medical training and the context in which their writing occurs: medical student, house officer, academic physician (full-time or part-time but working in an academic setting and involved in patient care, teaching, and/or research), and private practitioner (working in a nonacademic, clinical setting and not formally involved in teaching or research).

On the basis of the results of the literature survey and the interviews, a one-page questionnaire was designed. Because the term "medical records" includes types of writing relevant to different categories of physicians and trainees, depending on their level of experience, the term was broken down into particular types of inpatient records: write-up of history and physical examination, admitting note, progress note, discharge summary, operative report, and laboratory report. Some broader terms such as "teaching materials," "communications to patients," and "academic/administrative correspondence" were retained because a more specific breakdown would have made the questionnaire too long. Related types of writing were arranged together on the questionnaire.

In informal pilot testing of the questionnaire, copies were distributed to six individuals who had not been interviewed previously. Five of these were academic physicians with private practices and administrative experience and were from different departments; two of the five were departmental directors of medical education, and three were department chairmen. The sixth individual had a Ph.D. degree and was the director of an educational research center connected to a teaching hospital. After completing the questionnaire, the test group was asked whether the general format, directions, questions, and terms were clear; whether the categories of physicians and trainees and types of writing were valid; and whether all possible alternatives were given, items were ordered logically, and the questionnaire could be read easily and completed quickly. Final selection of the types of writing to be assessed resulted from this pilot testing of the questionnaire, and the term "other" was added to the types of writing to ensure that all types could be included.

On the revised questionnaire, respondents were asked, for each category of physician or trainee, to check the five types of writing most important for members of that category at any and all levels of experience within the category. "Importance" was defined on the questionnaire as referring to those types of writing that respondents felt were most needed to be taught if they were designing a writing

program for their curriculum. In this definition, it was also pointed out that the importance might or might not coincide with the frequency of that type of writing. Respondents also were asked to indicate for each type of writing whether their institution offered any course or program that specifically taught how to do that type of writing. Finally, respondents were asked whether any other systematic writing instruction was available at their institution, and space was provided for brief comments.

Questionnaires were mailed in the fall of 1985. A follow-up letter and another copy of the questionnaire were mailed six weeks later.

The questionnaires were sent to the 125 U.S. medical schools that had full institutional membership in the Association of American Medical Colleges (AAMC) as listed in the Spring 1985 *Group on Medical Education Membership Directory* (6). The questionnaires were sent in care of one of the school's designated members of the Group on Medical Education (GME) of the AAMC. GME members are appointed to this group by their respective deans, and the particular members surveyed had a defined interest in undergraduate or graduate medical education or both. At the time of the survey, a majority of the 125 selected GME members were associate deans for academic affairs and were identified as the "GME correspondent," the liaison between the correspondent's respective institution and the GME.

Results

Table 1 shows the 29 types of writing used on the questionnaire and the numbers of responding medical schools that offered formal instruction in these types of writing. The miscellaneous category "other" included creative and expressive writing and biographies for residency applications. These 29 types revealed a variety of contexts, purposes, audiences, uses, formats, media, and methods of transcription (longhand, shorthand, dictation, typing, word-processing). The audiences, for example, varied in size (one individual, small group, large group, mass audience), distance from the writer, level of education, and familiarity with the subject matter. A single type of writing such as a discharge summary might have multiple audiences and uses. Authorship itself might be single or multiple. Because their purposes differed, the types of writing also could exhibit different ways of thinking, organizational patterns, and stylistic features such as characteristic word choices and sentence structures.

A total of 100 (80 percent) of the 125 schools surveyed returned the questionnaire. Parts of all questionnaires were suitable for analysis; however, only 55 respondents complied with the direction that asked them to check only five types of writing for each of the four categories of physicians and trainees. Only these 55 questionnaires completed correctly were used to determine the major types of writing. Responses to each type of writing in each category of physicians and trainees were tallied. The five types of writing that received the highest number of responses in a particular category were considered to be the five most important kinds of writing for that category.

The five most important kinds of writing for the four categories are listed in Table 2. The highest responses regardless of category were considered to be the most important kinds of writing for all physicians and trainees. These were: write-ups of history and physical examinations, progress notes and discharge summaries (tied), peer-reviewed pub-

TABLE 1

Numbers of 100 Responding U.S. Medical Schools Offering Formal Instruction in 29 Types of Medical Writing, 1985

Types	No.
Write-up of history and physical examination	43
Admitting note	9
Progress note	18
Discharge summary	8
Letter to referring doctor	1
Operative report	1
Laboratory report	1
Outpatient records	5
Consultation report	3
Student evaluation	3
Faculty member evaluation	2
Teaching materials (lecture notes, outlines)	5
Research protocol	5
Grant proposal	12
Case history (not for publication)	1
Case report (for publication)	2
Chapter in textbook	0
Textbook	0
Peer-reviewed published paper (other than case report)	11
Abstract	4
Descriptive note (of new instrument, procedure for publication)	1
Literature review	5
Essay	3
Editorial	1
Book review	0
Letter to editor	1
Communications to patients (letter reporting test, examination results; instruction sheet)	2
Academic/Administrative correspondence (cover letter, memoranda, job/institutional applications)	2
Other (specify)	6

lished papers, and grant proposals (Table 2). Next most important were letters to referring doctor, outpatient records, consultation reports, admitting notes, student evaluations, and communications to patients, as shown in Table 3).

To determine the number of schools teaching particular types of writing, all the returned questionnaires were used. Table 1 shows, for each type of writing, how many of the 100 responding schools offered a course or program that specifically taught how to do that type. Forty-six schools did not offer formal instruction in any of the 29 types of writing listed on the questionnaire; at the other extreme, one school taught 14 of the 29 types.

The number of schools teaching the major types of writing for all physicians and trainees was examined. As shown in tables 1 and 3, 43 of the 100 responding schools offered a course or program that specifically taught physicians how to write up a history and physical examination; 18 taught how to write the progress note,

TABLE 2

Numbers of 55 Responding U.S. Medical Schools Ranking the Types of Medical Writing Considered Most Important for Each of Five Categories of Trainees and Physicians, 1985*

Category/Type	No.
Medical student	
Write-up of history and physical examination	54
Progress note	47
Admitting note	27
Outpatient records	26
Case history (not for publication)	23
House officer	
Discharge summary	47
Progress note	36
Write-up of history and physical examination	34
Admitting note	32
Outpatient records	24
Academic physician	
Peer-reviewed published paper	41
Grant proposal	38
Student evaluation	31
Consultation report	25
Teaching materials	25
Private practitioner	
Letter to referring doctor	37
Outpatient records	35
Consultation report	33
Discharge summary	30
Communications to patients	30
All	
Write-up of history and physical examination	54
Progress note	47
Discharge summary	47
Peer-reviewed published paper	41
Grant proposal	38

* Respondents were asked to rank from among 29 types of medical writing the five considered most important to be taught if they were designing a writing program for their curriculum. Of the 100 who returned the questionnaire, 55 complied with the instruction to rank only five types of writing.

12 the grant proposal, 11 the peer-reviewed published paper, and 8 the discharge summary.

On the whole, the five types of writing considered most important for all physicians and trainees (listed in tables 2 and 3) correlated well with the types most frequently taught, except for the discharge summary, as shown in Table 3. Twenty-seven of the 100 responding schools taught only one of the five types of writing considered important for all physicians and trainees, while 12 schools taught three or four. Concerning the types most important for medical students, 22 schools taught only one, while eight schools taught three to five. Of the types most important for house officers, 21 schools taught only one type, and 10 taught three to five. In regard to the types most important for academic physicians, 11 schools taught one, while three taught three of the five types. And for those types most important for private practitioners,

TABLE 3

Numbers of 55 Responding U.S. Medical Schools That Considered Each of 11 Types of Medical Writing Most Important To Teach and Numbers of 100 Responding Medical Schools That Offered Formal Instruction in the 11 Types of Medical Writing Most Frequently Taught, 1985

Considered Most Important To Teach			Taught by Greatest No. of Schools		
Rank Order	Type of Writing	No. ($n = 55$)	Rank Order	Type of Writing	No. ($n = 100$)
1.	Write-up of history and physical examination	54	1.	Write-up of history and physical examination	43
2.	Progress note	47	2.	Progress note	18
3.	Discharge summary	47	3.	Grant proposal	12
4.	Peer-reviewed published paper	41	4.	Peer-reviewed published paper	11
5.	Grant proposal	38	5.	Admitting note	9
6.	Letter to referring doctor	37	6.	Discharge summary	8
7.	Outpatient records	35	7.	Other	6
8.	Consultation report	33	8.	Outpatient records	5
9.	Admitting note	32	9.	Teaching materials	5
10.	Student evaluation	31	10.	Literature review	5
11.	Communications to patients	30	11.	Research protocol	5

eight schools taught only one type, while one taught three types.

Of the 100 responding schools, 48 taught none of the five major types of writing for all physicians and trainees, 56 taught none of the five major types for medical students, 57 taught none of the five major types for house officers, 79 taught none of the five major types for academic physicians, and 88 taught none of the five major types for private practitioners.

Only four of the 100 respondents indicated that their institution offered any systematic writing instruction other than formal teaching of specific types of texts. One of these four mentioned medical writing workshops for faculty members, while another noted writing seminars for students and fellows in biomedical research. Another indicated the availability of both a writing center that helped students improve their writing and a program that assigned writing fellows to individual medical faculty members who assigned papers in elective courses. No details regarding the writing center or type of writing fellow were given; however, the respondent explained that the fellows helped students prepare the assigned papers by critiquing their first drafts and then reviewing these drafts in individual conferences with the students, who later submitted the final draft, along with the critiqued draft, to the faculty member.

Discussion

According to the literature search reported here, the various kinds of physicians' writing have never been comprehensively classified and analyzed. Many types are taken for granted, although some, particularly the published kinds, are acknowledged as difficult for physician-writers. Even a brief analysis of the physician's possible writing tasks, however, reveals something of their diversity and complexity. Physicians and trainees with varying degrees of practice in both writing and medicine and under considerable time constraints compose texts in

school, hospital, clinic, laboratory, and office settings. Using a variety of methods of transcription, they write for varied purposes and audiences, and their texts reveal different methods of reasoning, patterns of organization, and stylistic features.

A majority of the respondents showed interest in the survey by their response rate of 80 percent. Although a number of the respondents had difficulty limiting their choices to five types of writing per category of physician and trainee, the survey showed that they generally agreed on the types of writing that are the most important for physicians to learn. While certain types of writing could be considered important to physicians and trainees as a whole, the respondents indicated that the writer's role (medical student, house officer, and so forth) appeared to determine the relative importance of the types of writing.

The respondents saw certain types of medical writing as most important to include in a medical curriculum; and their responses indicated that when medical schools formally taught specific types of writing, these were generally the types taught. However, the survey also showed that, with the exception of the history and physical examination write-up, the total number of responding schools actually doing formal teaching of specific types of writing was relatively small. According to the survey, a majority of the medical schools did not formally teach the major types of medical writing or offer any other kind of systematic writing instruction.

In answering the questions concerning formal teaching of specific types of texts and other available systematic writing instruction, several respondents spontaneously commented on the availability of informal instruction within the medical curriculum. For example, they mentioned brief instruction by clerkship directors as students came on rotation or informal critiques of individual writings by faculty members and residents. Hypothetically, informal teaching would also include the various self-teaching strategies medical students and house officers engage in, such as imitating another physician's writing or following some sort of standardized model (7). A good deal of informal teaching thus may be present in medical schools; however, the questionnaire did not measure its frequency.

Even if school instruction, formal or informal, is available, it might not address important questions such as whether or not information is appropriately selected and effectively organized in the light of the purposes, audiences, and uses for the writing. For example, the nature of even the most fundamental type—the write-up of a patient's medical history and physical examination—varies as its purpose, audience, and use change from one kind of writer and context to another. The lengthy history and physical examination write-ups prepared by students for their teachers must show what the students have learned. Later in their training, the students must produce more concise write-ups that communicate precisely relevant information to multiple audiences outside of the school setting (7).

Skill in writing the first, lengthier kind of write-up may not ensure a student's success in the latter task because the ability to write effectively for one purpose and audience does not necessarily imply the ability to write effectively for another (8). As Sox and colleagues (4) noted, medical students "need help in making the transition from an excessively detailed written medical record and one that communicates succinctly what is needed to care for the patient." However, adequate instruction may not be available either in medical schools or in the literature that

students or practicing professionals might refer to.

The literature on medical writing discusses the scientific paper, the grant proposal, and, to some extent, the history and physical examination portion of the medical record. Few articles or books treat other important types of writing, such as the admitting or progress note, letter to referring doctor, or discharge summary. Also, discussions in the literature tend to emphasize writing for publication and to offer editorial recipes for effective writing or advice on final published products rather than clear, detailed explanations of the processes involved in composing them and of key principles, such as their purposes, uses, and audiences.

The present authors consider that the lack of formal school and textbook instruction in the major types of medical writing points to an important but little explored area for future teaching. The present survey supports the need to develop curricula in medical schools that address the systematic and effective teaching of medical writing.

References

1. Stillé, A., Smith, Jr., F. G., Yardley, T. H., and Johnson, J. B. Report of the Committee on Medical Literature. *Trans. Amer. Med. Assoc.*, **3**:153–187, 1850.
2. Medical Literature and Medical Writing. *J.A.M.A.*, **35**:626–627, 1900.
3. Woods, D. Good English Is Good Medicine. *Can. Med. Assoc. J.*, **15**:624–629, 1981.
4. Sox, Jr., H. C., Morgan, Jr., W. L., Neufeld, V. R., Sheldon, G. F., and Tonesk, X. Subgroup Report on Clinical Skills. In Physicians for the Twenty-First Century/Report of the Project Panel on the General Professional Education of the Physician and College Preparation for Medicine. *J. Med. Educ.*, **59**:139–147, Part 2, November 1984.
5. Bjork, R. E., and Oye, R. K. Writing Courses in American Medical Schools. *J. Med. Educ.*, **58**:112–116, 1983.
6. *Group on Medical Education Membership Directory.* Washington, D.C.: Association of American Medical Colleges, 1985.
7. Pagano, M. P., and Mair, D. Writing Medical Records. *J. Tech. Writ. and Com.*, **16**:331–341, 1986.
8. Britton, J. L., Burgess, T., Martin, N., McLeod, A., and Rosen, H. *The Development of Writing Abilities (11–18).* London: Macmillan Education, 1975.